

Toxicoinformatics at the NCTR

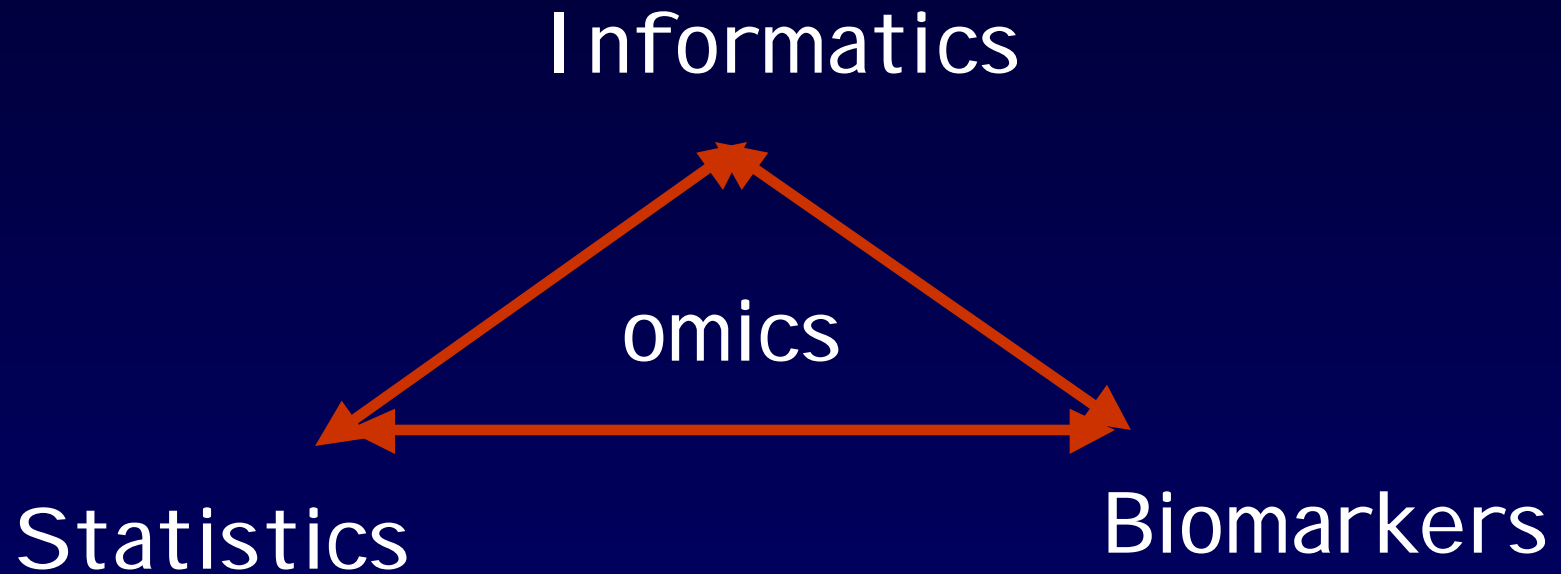
Daniel A. Casciano, Ph.D.

National Center for Toxicological Research
Food and Drug Administration

NCTR Strategy

- Develop a validated gene expression data base in surrogate organisms
- Develop a validated gene expression data base in humans (long-term goal)
- Develop a proteomic capacity
- Develop a metabonomic capacity
- Utilize existing expertise in statistics and bioinformatics to support these initiatives

Systemic Approach to Toxicogenomics



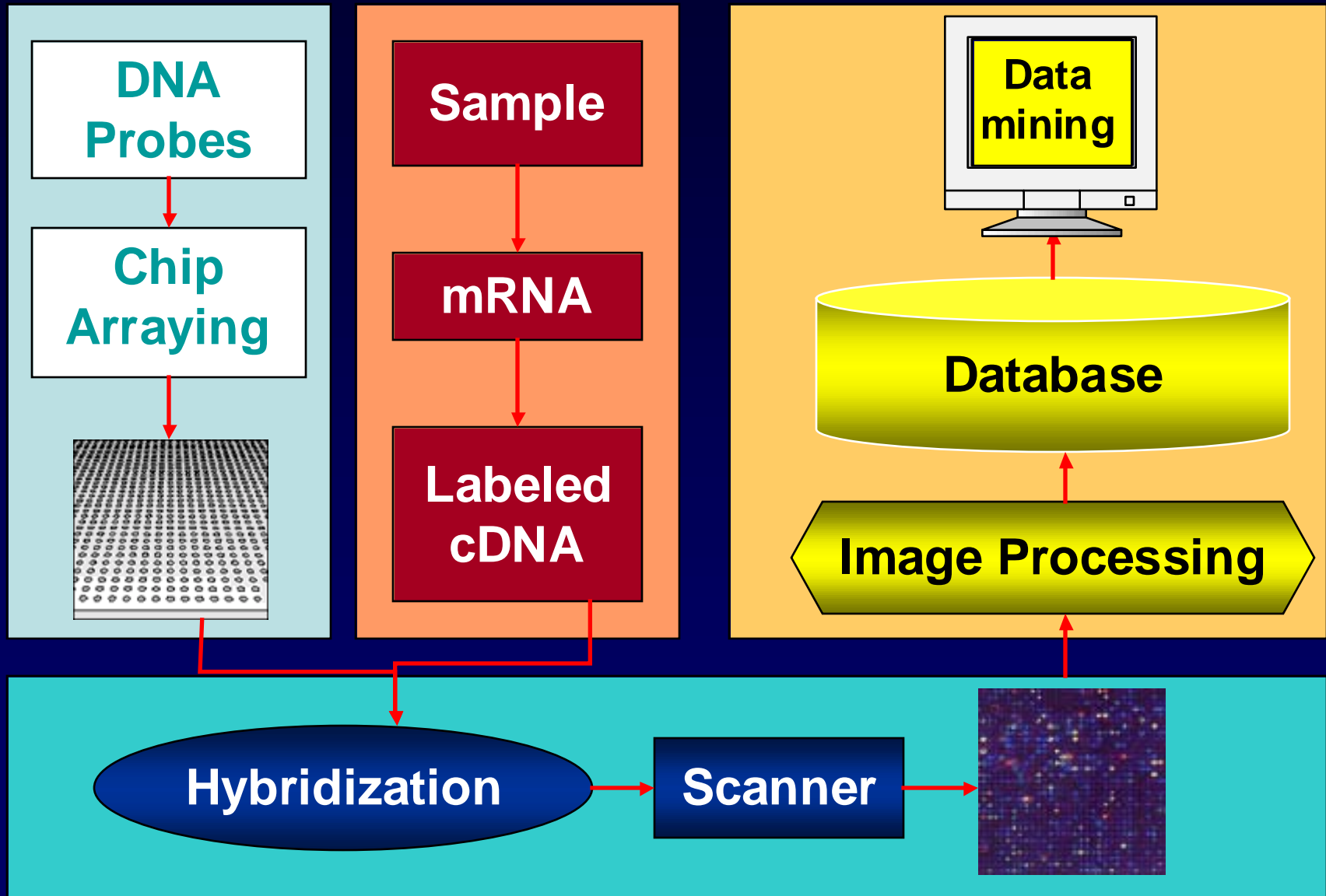
Centers of Excellence

- The Functional Genomics Center
- The Structural Genomics Center
- The Toxicoinformatics Center
- The Hepatotoxicity Center
- The Phototoxicity Center

Functional Genomics

- the study of gene function on a whole or partial genome scale that may include the study of gene expression using DNA microarray

Components of Microarray Experiment



Functional Genomics Center

- Purchased necessary equipment to print and scan slides
- Developed a core center to standardize molecular, analytical, and informatic tools
- Several protocols from various NCTR divisions in various stages of preparation

Functional Genomics Center

- We have purchased a collection of rat, mouse and human genes.
 - 4000 rat genes
 - 5000 mouse genes
 - 8300 human genes
- Cost per slide is ~10% of commercially prepared slides.

Types of Genes

- Toxicology, drug metabolism
- DNA metabolism, replication, repair
- Cell cycle, apoptosis
- Oncogenes, tumor suppressors, receptors
- Intermediary metabolism, hormones
- Stress response
- Inter- and Intra-cellular communication
- Transcription factors
- etc., etc., etc.,....

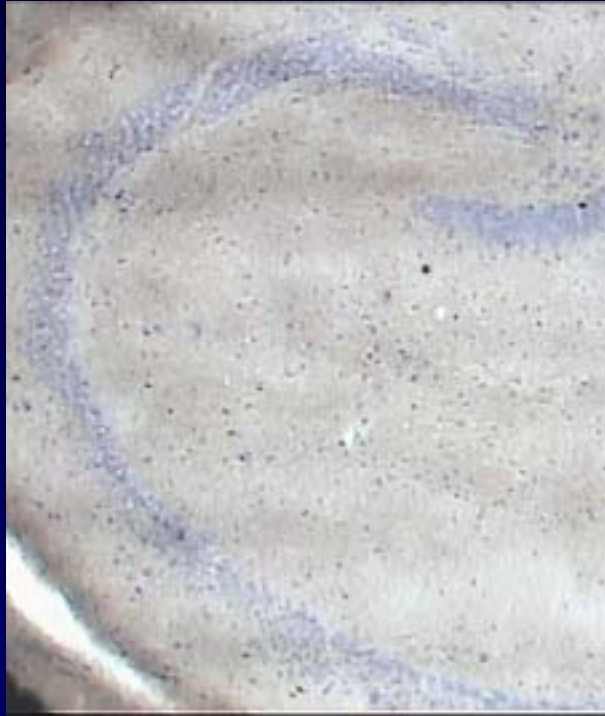
Status

- Standardizing techniques
- Developing a mito chip
- Defining control gene expression patterns in aging Fisher 344 rats
- Defining circadian gene expression in Fisher 344 rats
- Developing collaborations within and without the NCTR

Microarray Improvements

- Reduction in variability
- Stronger signals – better labeling of cDNA
- Reduced autofluorescence – BSA treatment, sodium borohydride
- Increased uniformity across slide – lifter slips instead of coverslips
- Coefficients of variability are now 12-15%

Laser Capture Microdissection CA Region of Hippocampus



Before



After



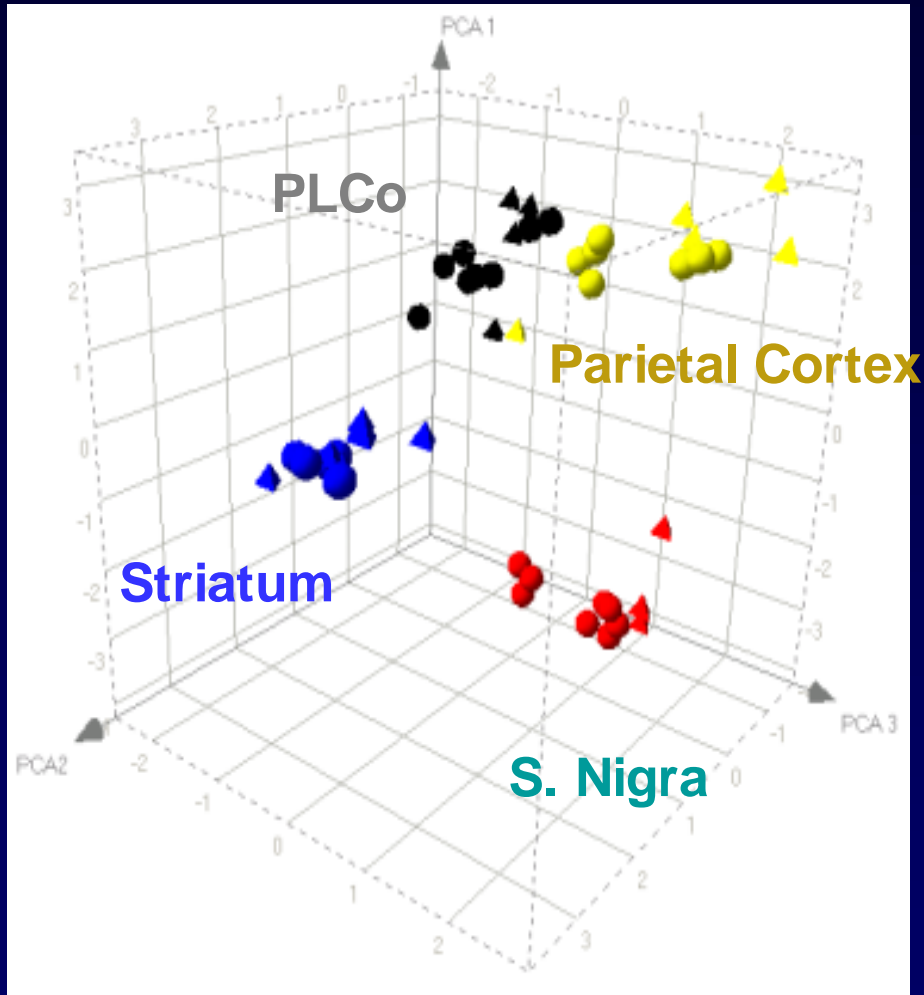
Cells on cap

Gene Expression Pattern in 4 Brain Regions

Objective – Relationship of regional differences with gene expression pattern

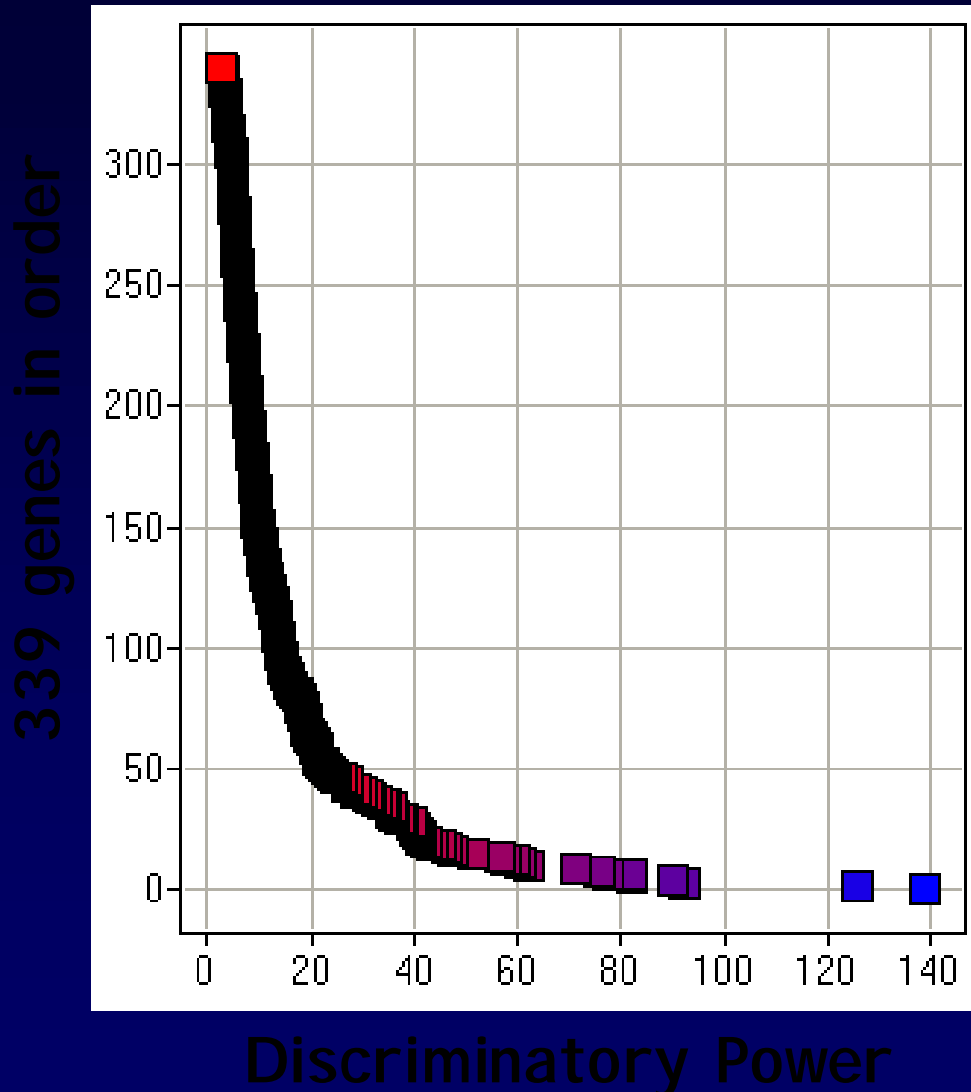
- Filter array – Clontech Atlas 1.2K rat array
- 29 mRNA samples from 4 brain regions
 - S.Nigra (8 samples)
 - Striatum (6 samples)
 - Parietal Cortex (PC, 7 samples)
 - Posteriolateral cortical amygdaloid nucleus (PLCo: 8 samples)
- Clustering: PCA and 2-way HCA

Principle Component Analysis (PCA) Model



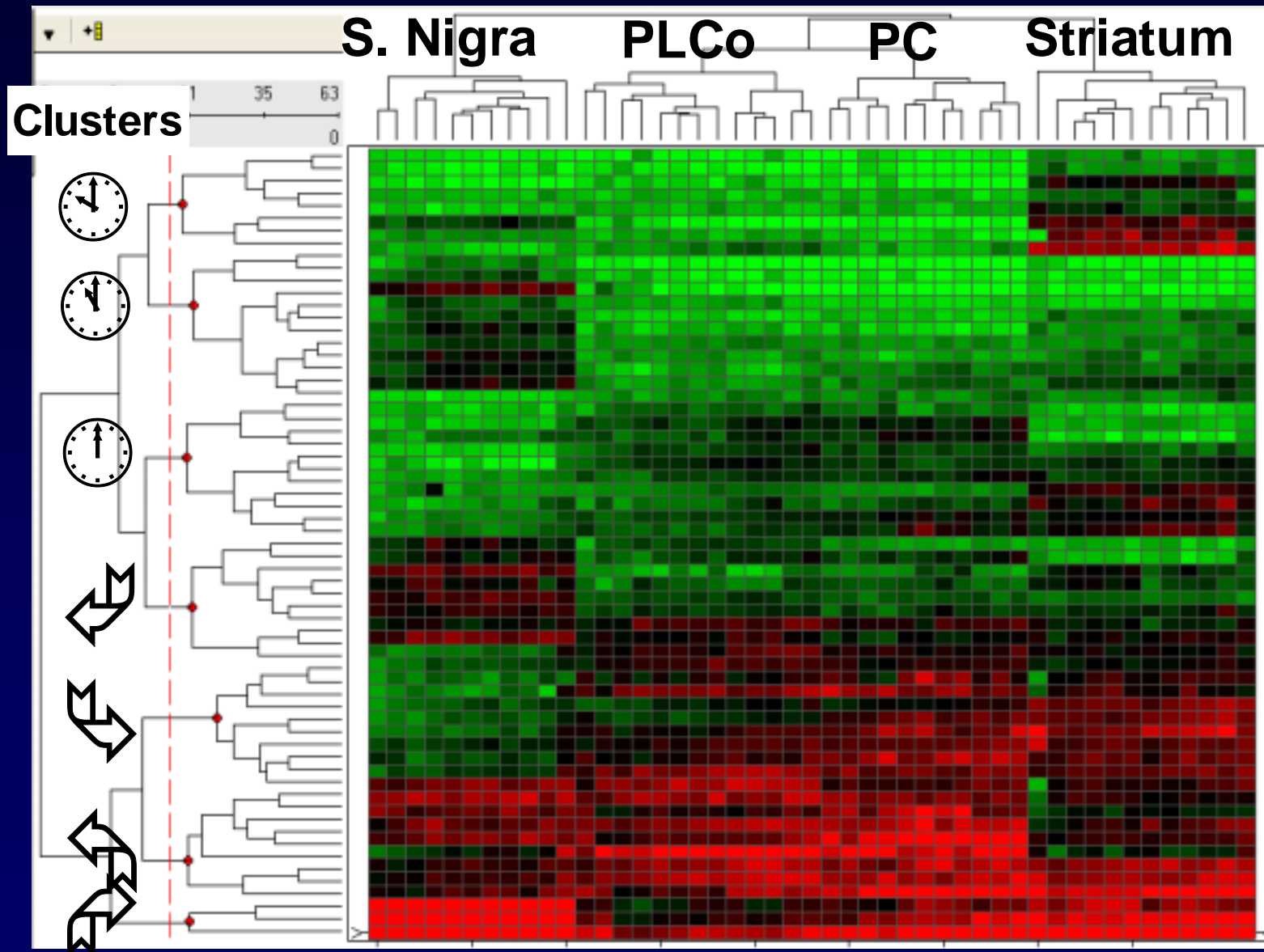
- 29 samples for the model (circle)
- 18 new samples are tested, and their identities are not revealed (triangle).
- Conclusion: PCA using 339 differentially expressed genes can efficiently separate the 4 brain regions.

SIMCA – Identification of discriminatory genes



- Rank order 339 genes based on their discriminatory power (DP) to distinguish 4 regions using SIMCA
- The majority of genes has the DP < 20
- 63 genes > 20

2-Way Clustering

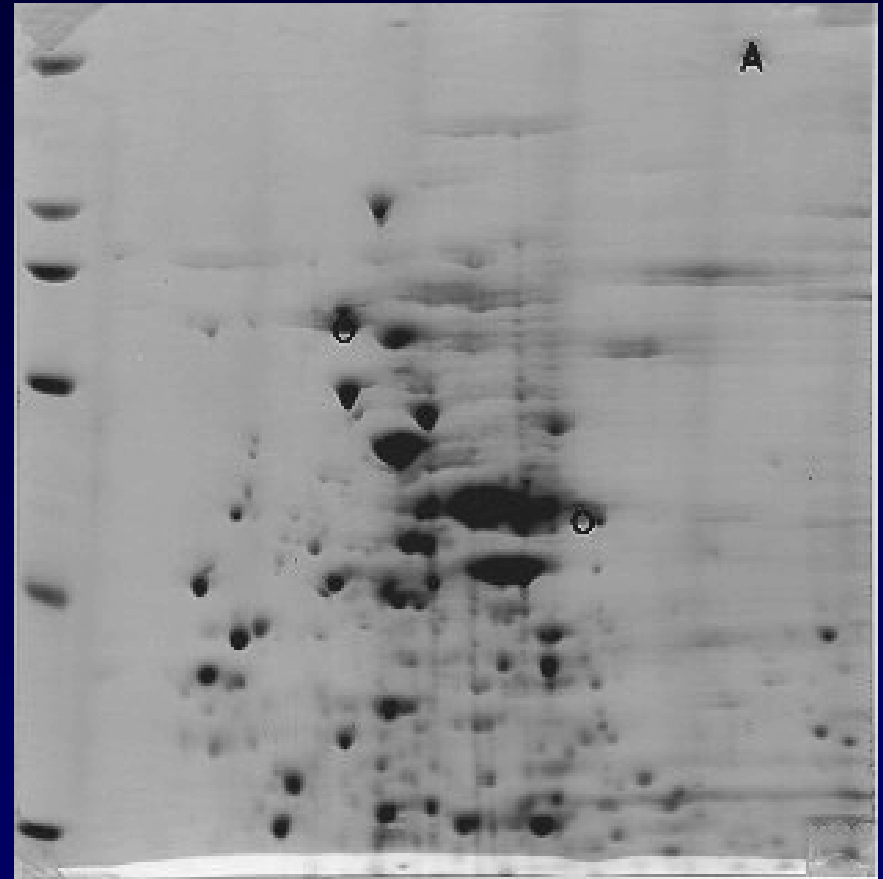


Proteomics

Hepatotoxicity–idiosyncratic
drug reactions

PC-based
visualization/quantitative
analysis of 2-D gels

Robotics: sample retrieval
and preparation.



2D-Gel Electrophoresis

Proteomics

Chemistry

Biomonitoring metabolites, biomarkers
(Micromass Ultima Quattro Triple Quad)

Detection large intact biomolecules
(Seldi-TOF; MALDI's TOF; MAB TOF)

Protein Sequence Identification and
fragmentation patterns (QTOF)

Rapid identification of food-borne
pathogens and biological weapons of
mass destruction

Status

- Developed an infrastructure
- Recruited analytical and biological staff
- Defining the mitoproteome of Fischer 344 rat
- Collaborations within and without the NCTR

Metabonomics

- the quantitative measurement of the dynamic multiparametric metabolic response of living systems to pathophysiological stimuli or genetic modification

Metabonomics

- Purchased 600 mHZ NMR
- Concept papers approved for protocol development

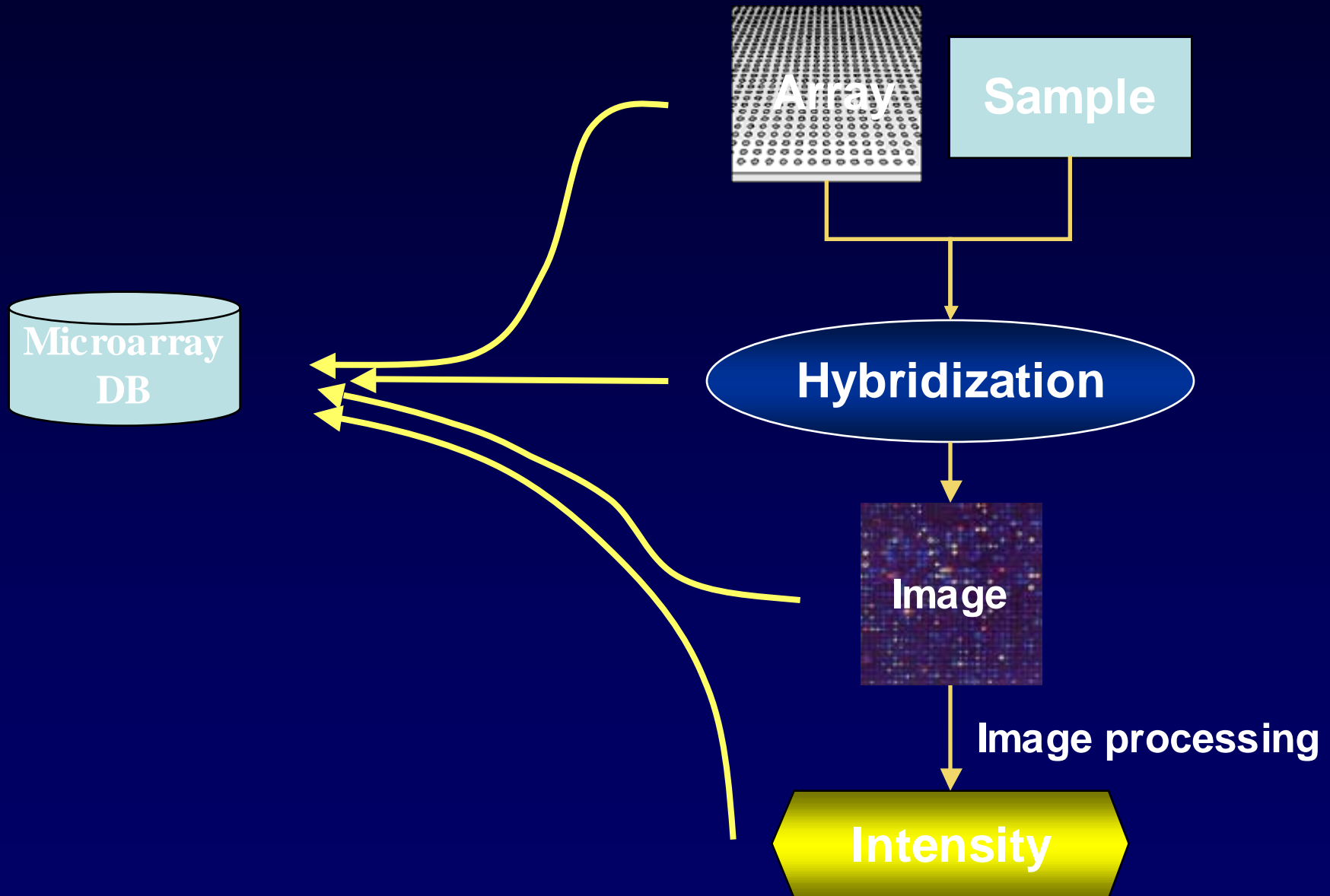
Toxicoinformatics

- Provide effective software infrastructure and analysis capabilities to address the following challenges:
 - managing and integrating different data sources
 - analyzing and visualizing these data to formulate new regulatory standards

Omics Research at NCTR

- Traditional toxicological research
 - Animal models
 - Conventional methods (*in vitro* and *in vivo* assays)
- Omics research – genomics, proteomics and metabonomics
 - High throughput technology is a driving force for omics research
 - Characteristics - large volume of data
- Examples:
 - Genomics: Microarray
 - Proteomics: 2D-gel/MS and SELDI-TOF
 - Metabonomics: NMR

ArrayTrack: DB-Lib-Tool Structure



MicroarrayDB – Validated Database

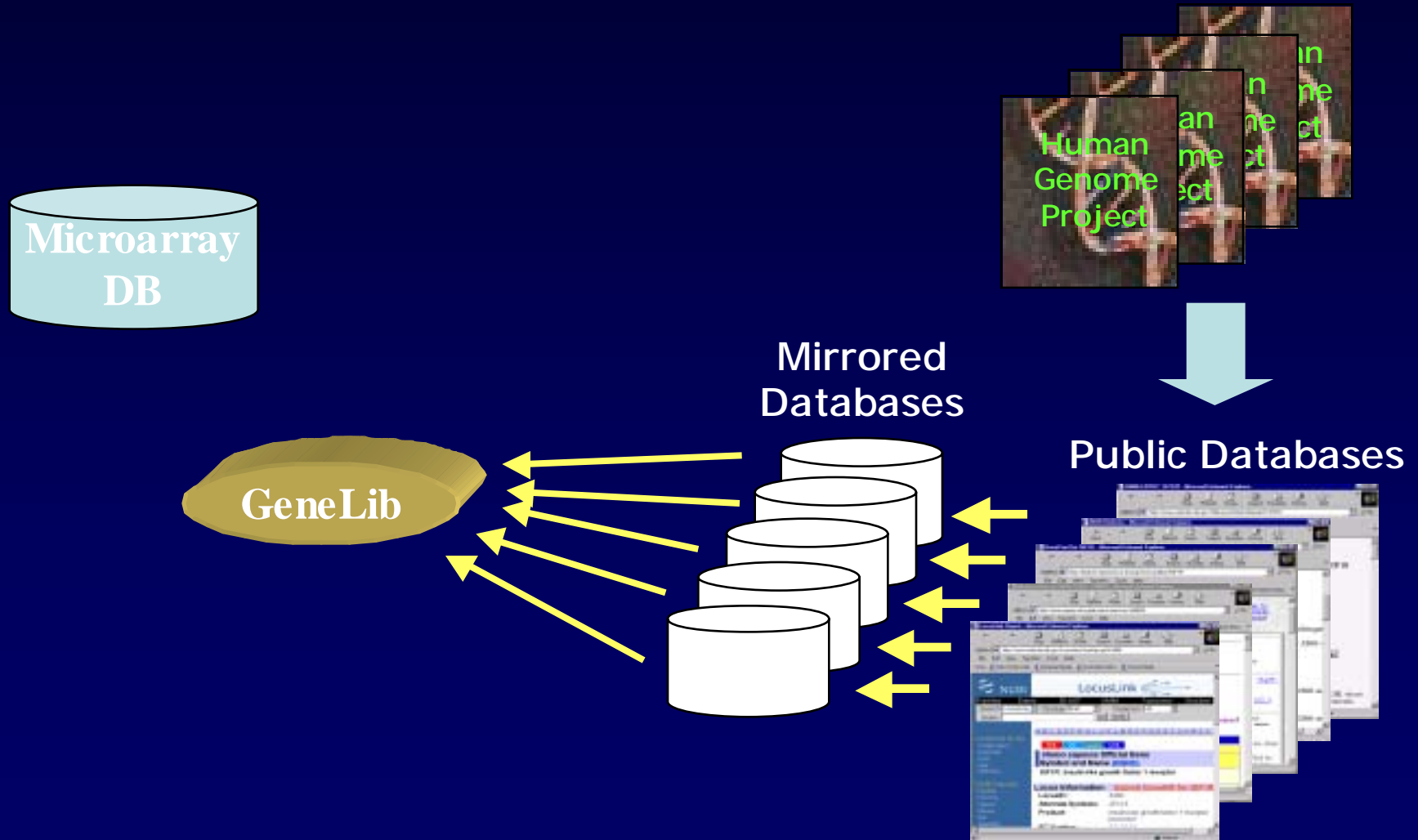
Objective: Data entered only if they meet prescribed standards for completeness, accuracy (QA/QC) and ontology

- Establish standardized experimental protocols (hybridization, labeling and RNA extraction) at NCTR
- Data stored are defined by the MIAME (Minimum Information About a Microarray Experiment) standard
 - Results can be interpreted
 - Experiment can be independently verified
- Controlled vocabulary and ontology defined by MGED ontology working group

MicroarrayDB – Validated Database (cont.)

- Specify toxicology-relevant information
 - MIAME/Tox (collaborating effort from ILSI, EBI and NCT)
 - Developed based on MIAME 1.1
 - Define minimum (additional) info for toxicogenomics research
 - ArrayTrack is MIAME/Tox compliant
- Security
 - Data are protected at the individual experiment level
 - Owner of the data defines read and/or write protection of the data

ArrayTrack: DB-Lib-Tool Structure



Specify ID Type:

- ☐ GeneSymbol
☐ GeneBankAcc
☐ UnigeneID
☐ LocusID
☐ GEN_ID_MFT



GEN

Enter Gene

AADAC

ACACB

ACACA

ACAT2

ACAA1

ACAT1

ACAT2

ACAT1

ACAA1

ACADM

ACADL

ACADS

ACAT2

ACAT1

ACADVL

ACAA1

ACADS

ABAT

ABAT

AARS

ABP1

ACADS

ACADM

ACAA1

ACAT2

ACAT1

ABP1

ABP1

ABP1

ACADVL

ACAA1

ABP1

ACADS

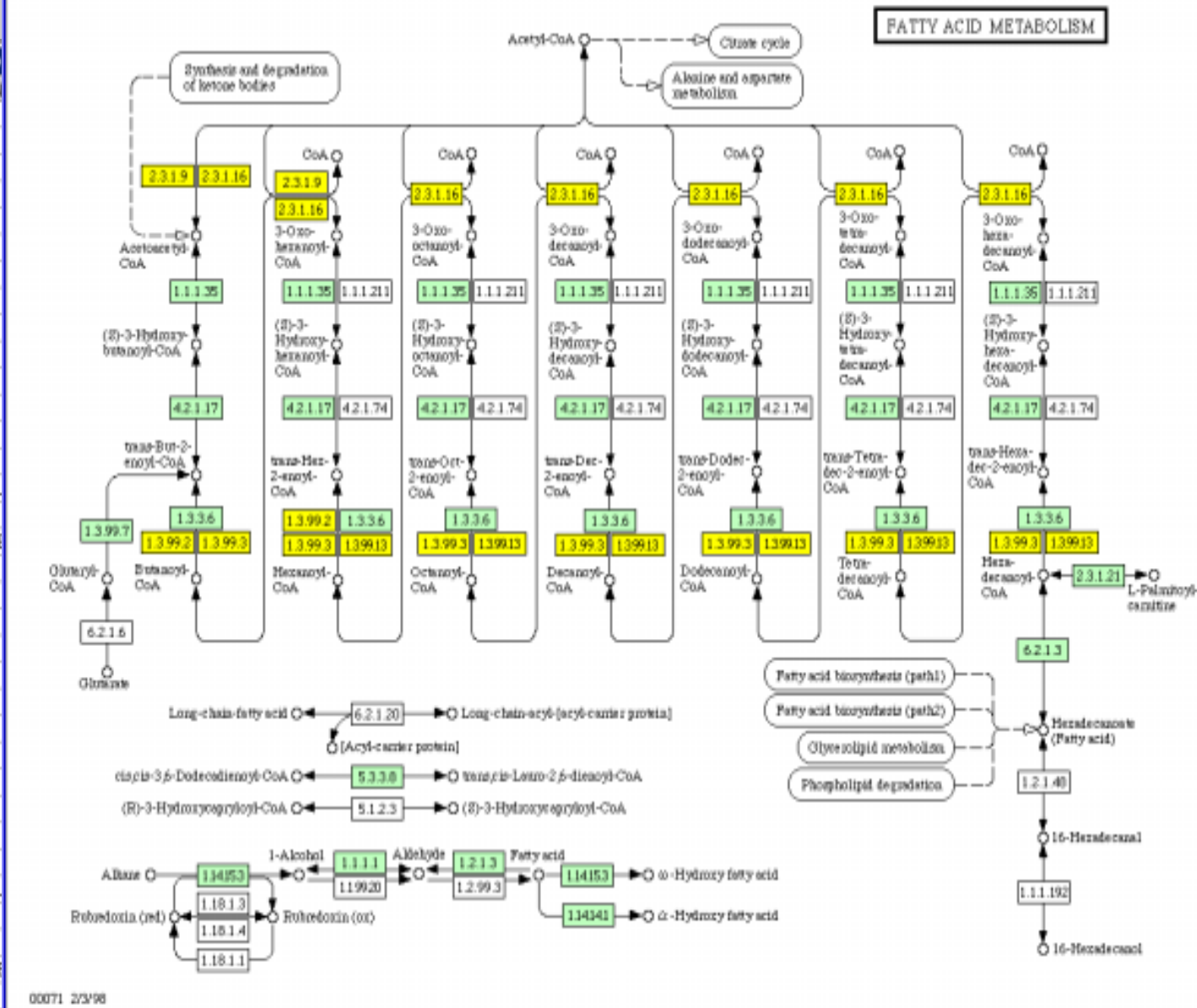
ACAT2

AANAT

Search



Clear



eneLib

Gene Search

More Info ... Select one -- Go

Link To ... Select one -- Go

View Pathway Customize Table Export Table

Specify ID Type:

- ☐ GeneSymbol
- ☒ GeneBankAcc
- ☐ UnigeneID
- ☐ LocusID
- ☐ GEN_ID_MFR

Enter Gene IDs

Search ☐ within result

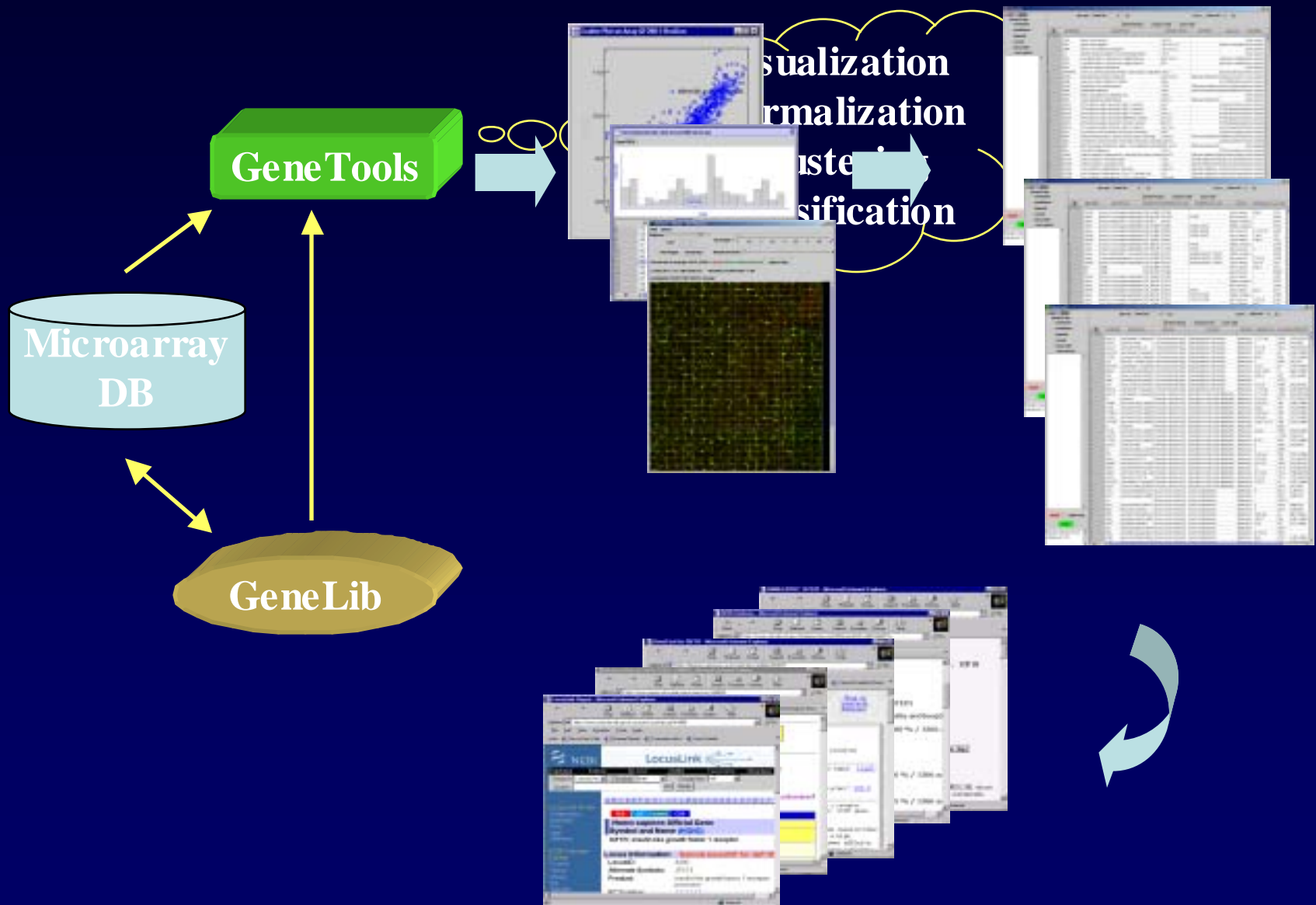
Clear

Message: total gene records in library are 159,271.
Missing number: 1. Missing

GENEID	GENENAME	DESCRIPTION	CHROMLOCATION	PATHWAYS
1	PRNP	prion protein (p27-30) (Creutzfeld-Jakob disease, Gerstmann-Strausler-Scheutler)	20pter-p12	
2	ADH1B	alcohol dehydrogenase 1B (class I), beta polypeptide	4q21-q23	Tyrosine metabolism;Fatty acid metabolism;Bile acid biosynthesis
3	ELA3A	elastase 3A, pancreatic (protease E)	1p36.12	
4	TYMS	thymidylate synthetase	18p11.32	Folate biosynthesis;Pyrimidine metabolism;One carbon pool
5	CD7	CD7 antigen (p41)	17q25.2-q25.3	
6	PSMB1	proteasome (prosome, macropain) subunit, beta type, 1	6q27	Proteasome
7	PSMA3	proteasome (prosome, macropain) subunit, alpha type, 3	14q23	Proteasome
8	MARCKB	myristoylated alanine-rich protein kinase C substrate	6q22.2	
9	ODDH	oxoglutarate (alpha-ketoglutarate) dehydrogenase (isozyme)	7p14-p13	Lysine degradation;Tryptophan metabolism;Citrate cycle (TCA)
10	PMP22	peripheral myelin protein 22	17p12-p11.2	
11	EIF4G1	eukaryotic translation initiation factor 4 gamma, 1	3q27-qter	
12	ETV4	ets variant gene 4 (E1A enhancer binding protein, E1AF)	17q21	
13	GL01	glyoxalase I	6p21.3-p21.1	Pyruvate metabolism
14	CCT8	chaperonin containing TCP1, subunit 8 (theta)	21q22.11	
15	KIAA0008	Drosophila discs large-1 tumor suppressor-like	14q22.1	
16	CHPFR	likely ortholog of chicken chondrocyte protein with a poly-proline region	8q12.1	
17	TCOM20-PEND	translocase of outer mitochondrial membrane 20 (yeast homolog)	1q42	
18	KIAA0020	KIAA0020 gene product	9p24.2	
19	KIAA0101	KIAA0101 gene product	15q22.1	
20	AOP2	anti-oxidant protein 2 (non-selenium glutathione peroxidase, acidic calcium)	1q23.3	Methane metabolism;Phenylalanine metabolism;Flavonoids, s
21	KIAA0107	KIAA0107 gene product	3p14.3	Proteasome
22	ADAM9	a disintegrin and metalloproteinase domain 9 (metrin gamma)	8p11.21	
23	KNSL2	kinesin-like 2	6p21.3	
24	PTDSS1	phosphatidylserine synthase 1	6q22	Glycerolipid metabolism;Aminophosphonate metabolism;Eph
25	MROX	MORF-related gene X	10q22	
26	PTPRH	protein tyrosine phosphatase, receptor type, H	19q13.4	
27	ACAA2	acetyl-Coenzyme A acyltransferase 2 (mitochondrial 3-oxoacyl-Coenzyme	18q21.1	Fatty acid metabolism;Bile acid biosynthesis;Phenylalanine me
28	NUDT1	nucleoside diphosphate linked moiety X-type motif 1	7p22	
29	TAGLN2	transgelin 2	1q21-q25	
30	NOLC1	nucleolar and coiled-body phosphoprotein 1	10q24.32	
31	LARS2	leucyl-tRNA synthetase, mitochondrial	3p21.3	Aminoacyl-tRNA biosynthesis;Valine, leucine and isoleucine bi
32	RRS1	homolog of yeast ribosome biogenesis regulatory protein RRS1	8q12.3	

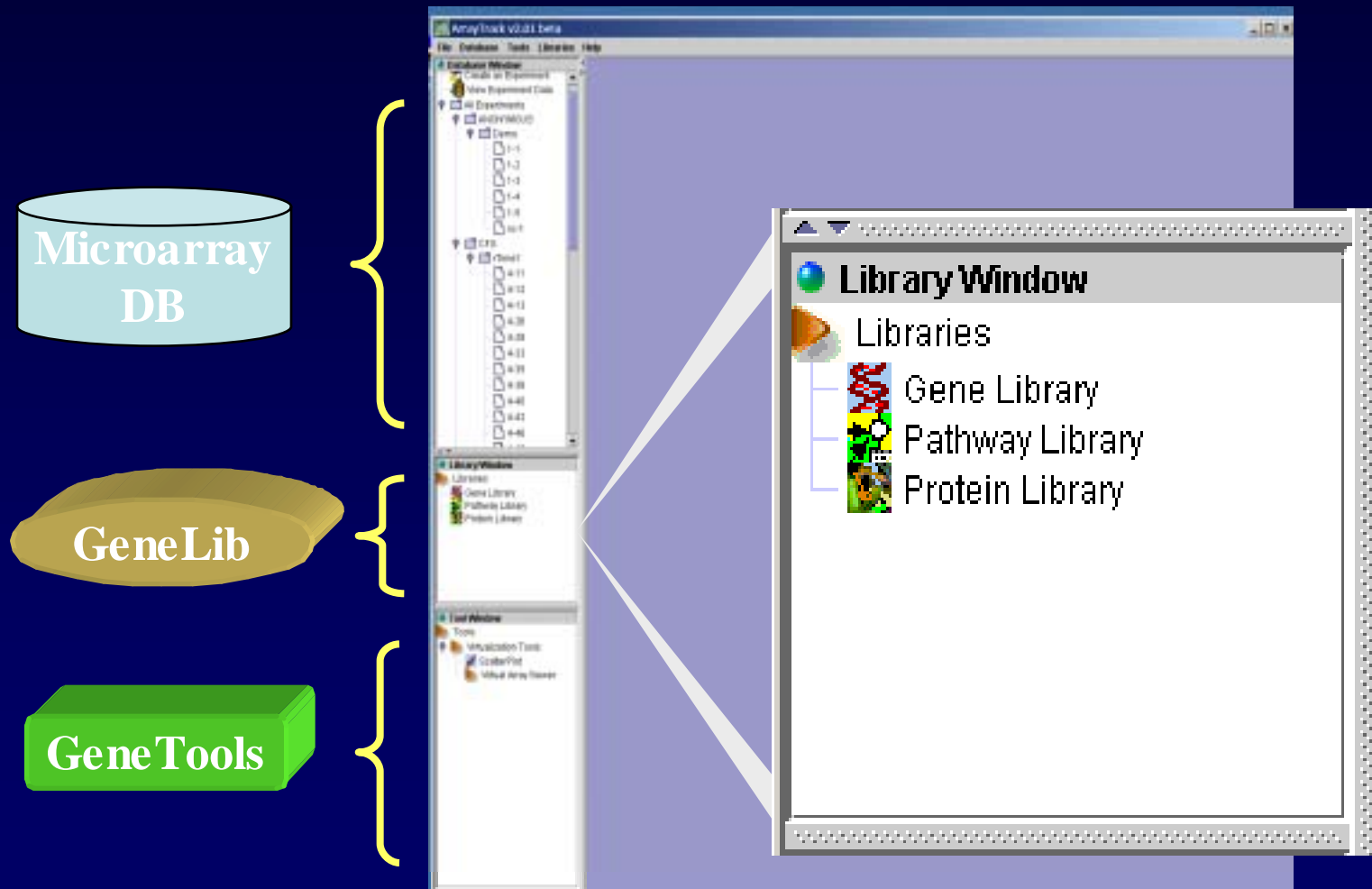
Import a list of genes from other software (ex. Excel)

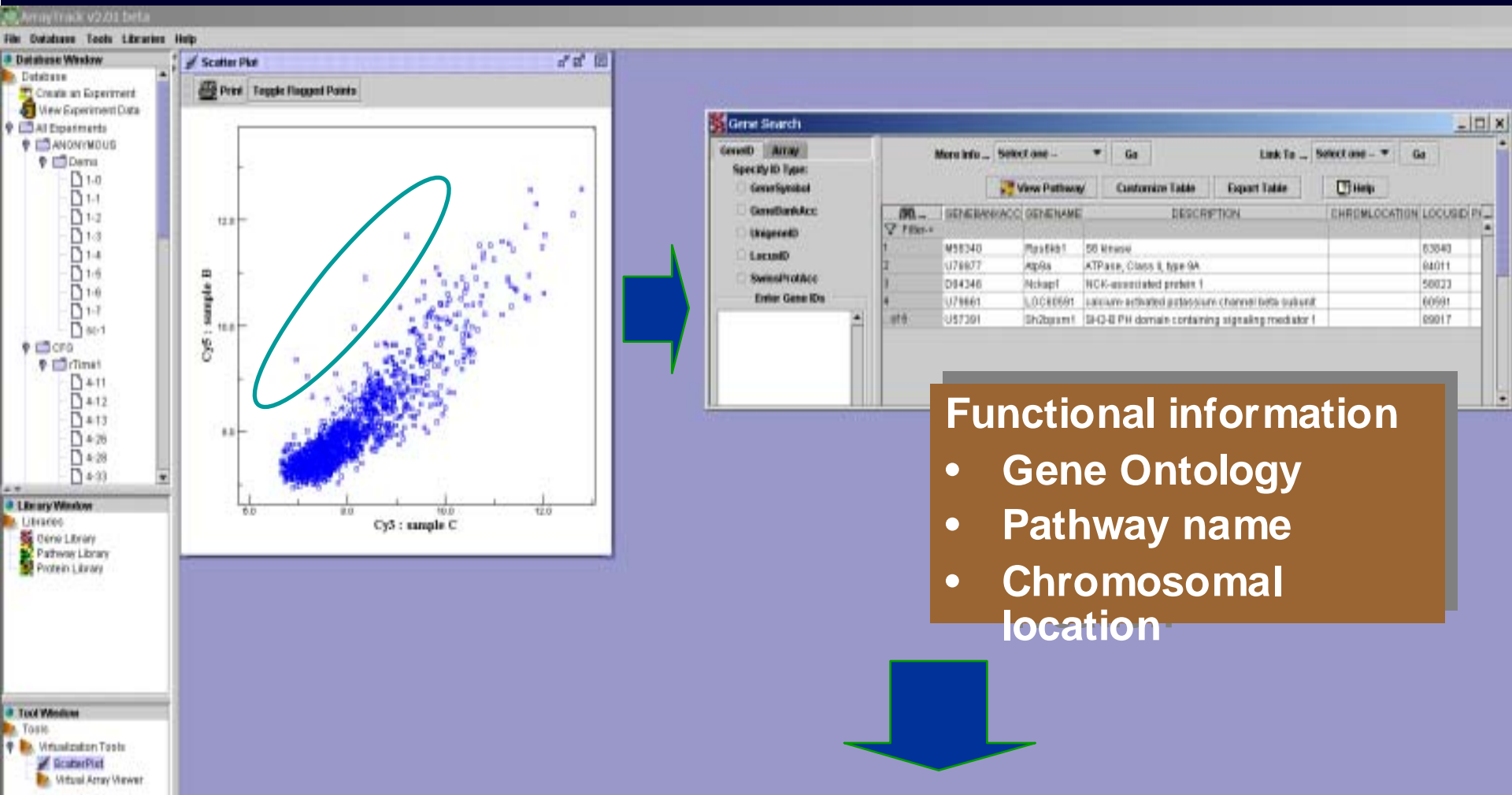
ArrayTrack: DB-Lib-Tool Structure



ArrayTrack

<http://weblaunch.nctr.fda/jnlp/arraytrack>





Functional information

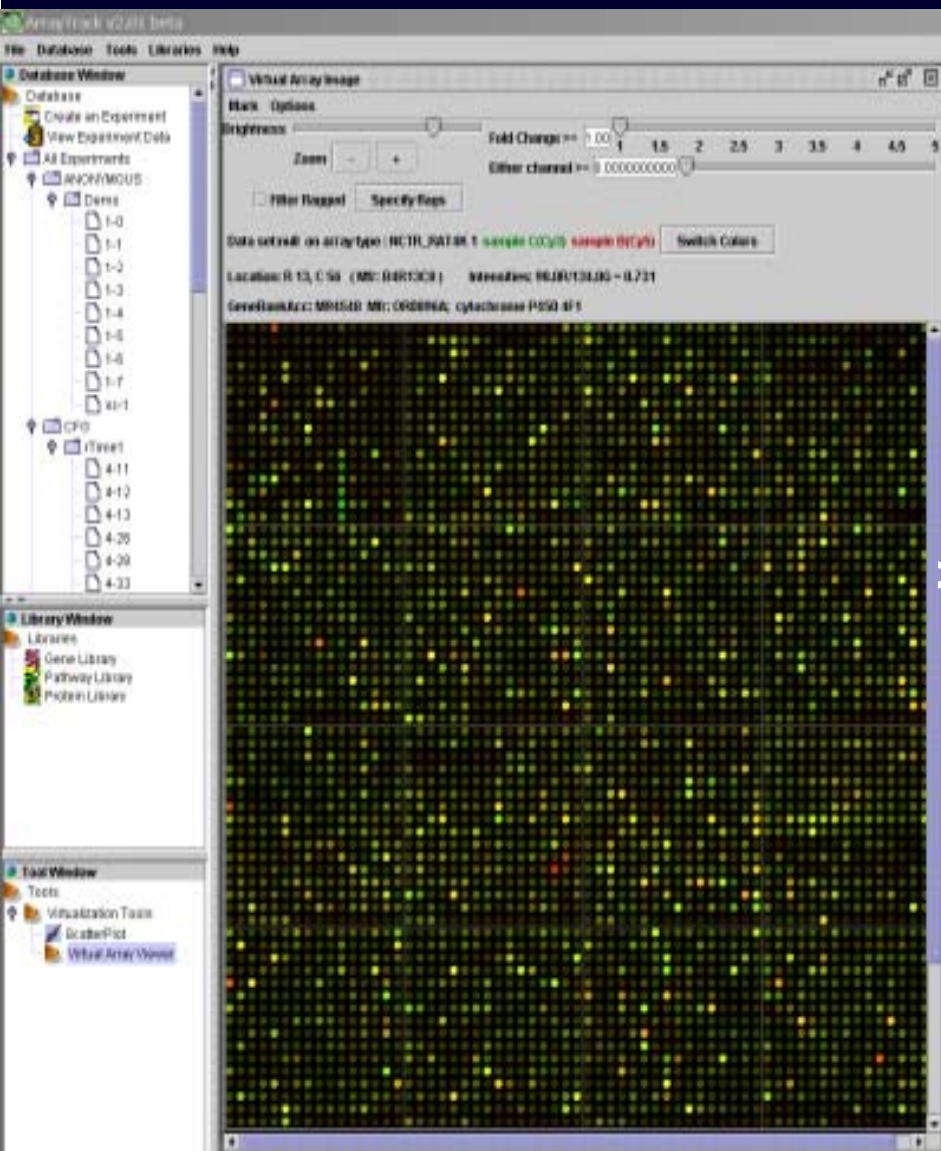
- Gene Ontology
- Pathway name
- Chromosomal location

Pathway (from Sapiens only)

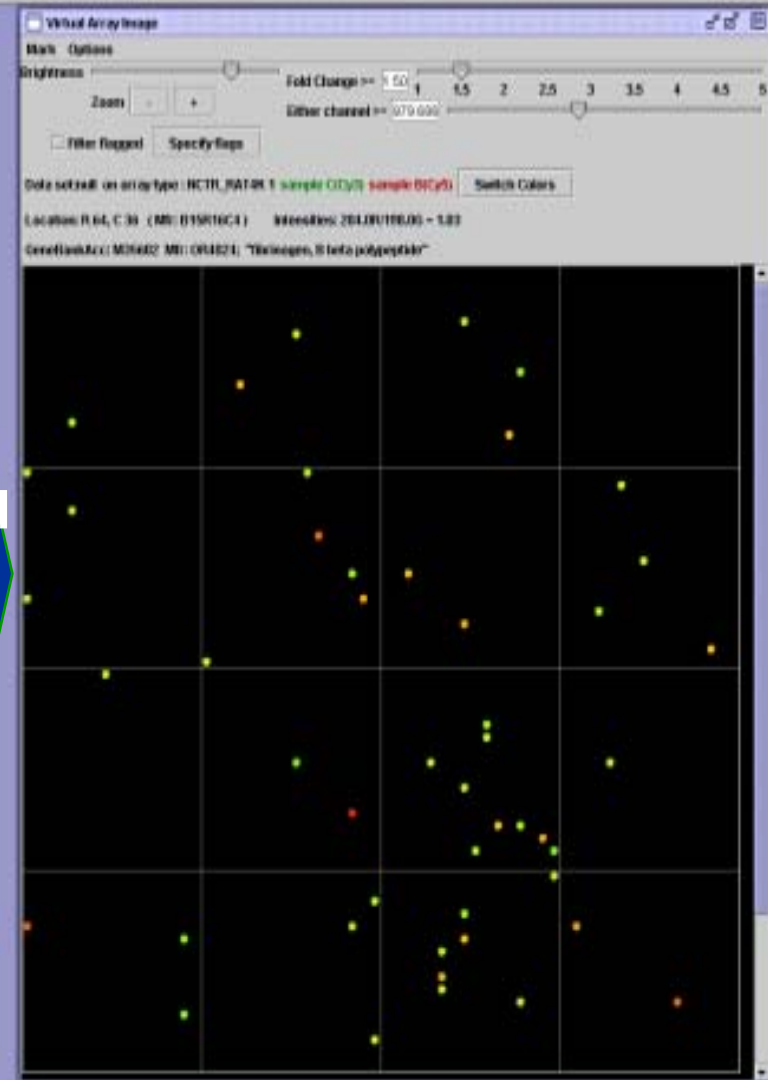
Export pathway data

Genes	Map	Category	Compounds	NCCT/Toxicology
RP06-B1	Starch and sucrose metabolism (hsa00500)	Metabolism of Complex Carbohydrates/Metabolic pathway	53	1
RP06-B1	Inositol phosphate metabolism (hsa00562)	Metabolism of Complex Lipids/Metabolic pathway	26	0
RP06-B1	Sphingolipid metabolism (hsa00603)	Metabolism of Complex Lipids/Metabolic pathway	21	0
RP06-B1	Benzene degradation via Catechol (hsa00362)	Degradation of Xenobiotics/Metabolic pathway	46	5
RP06-B1	Burkholderia and related genera metabolism (hsa00793)	Metabolism of Coliforms and Relatives/Metabolic pathway	23	2

Total submitted genes: 1 genes found, -1 not found, Total 5 pathway maps

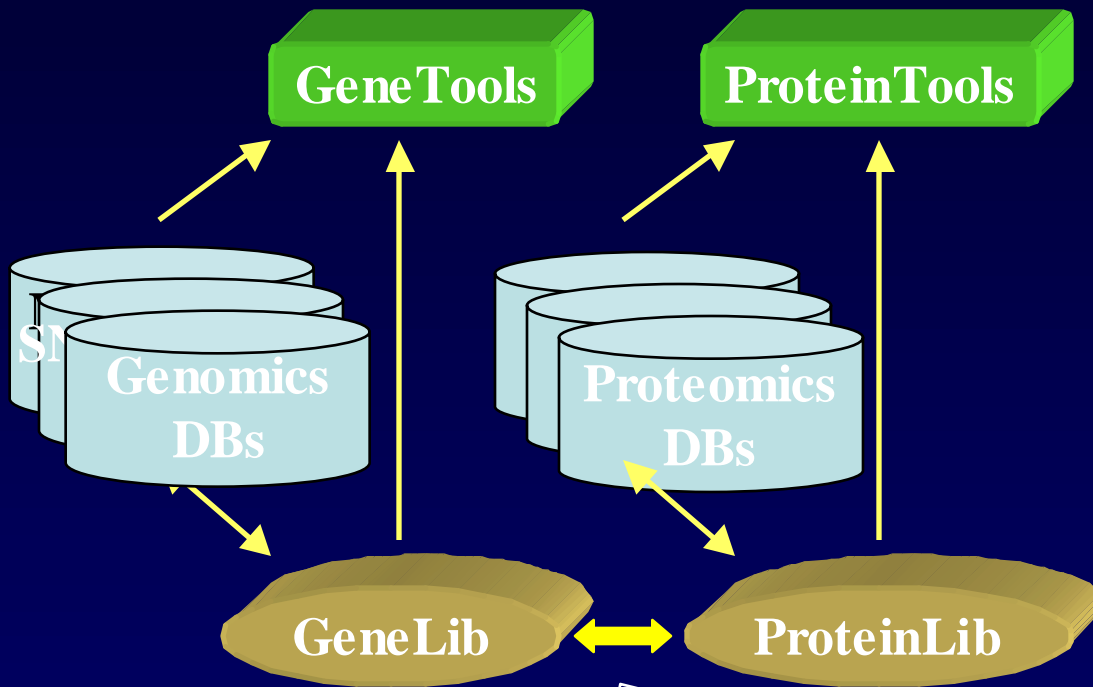


>1.5 fold
remove
low int.



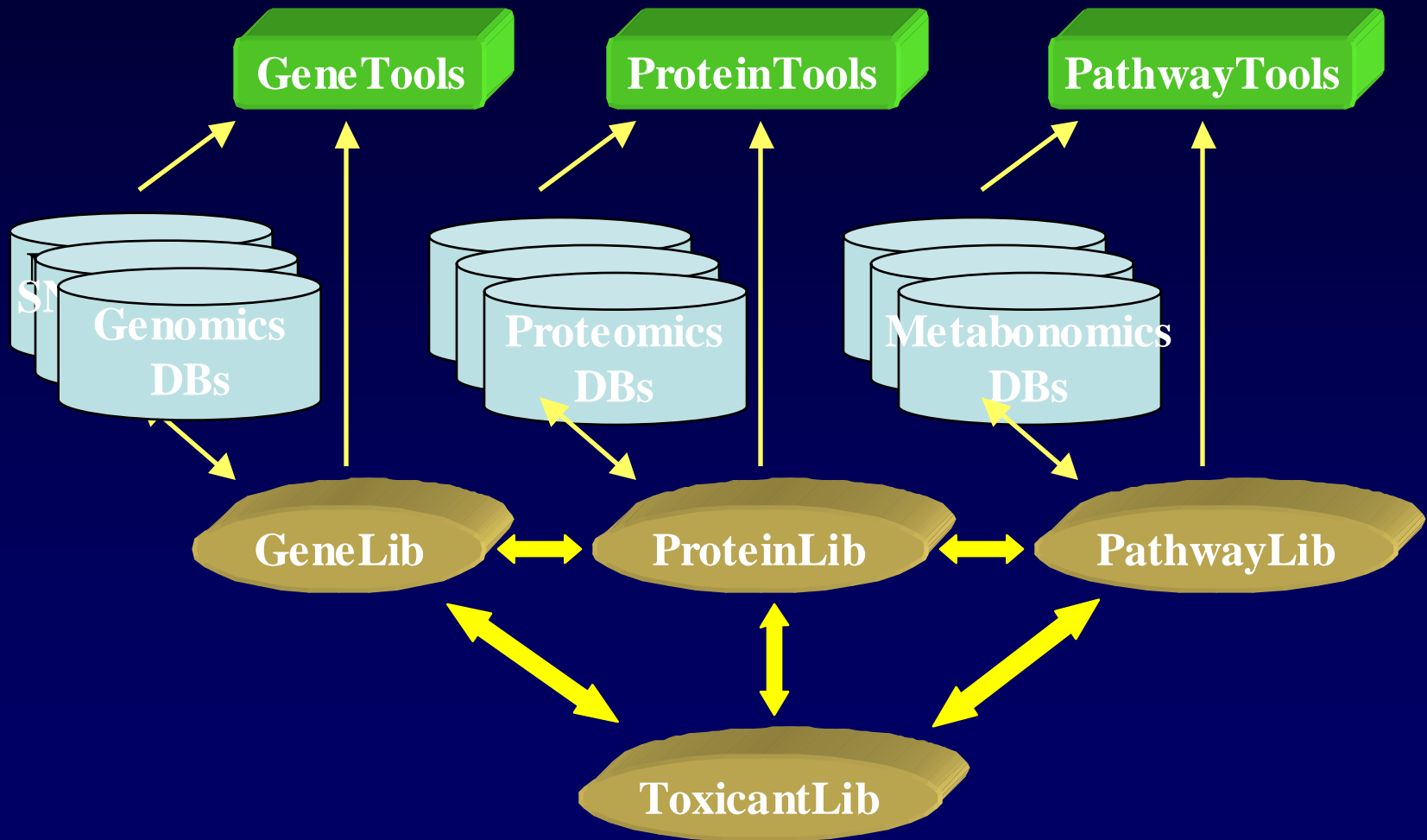
Link to
GeneLib

Proteomics Data

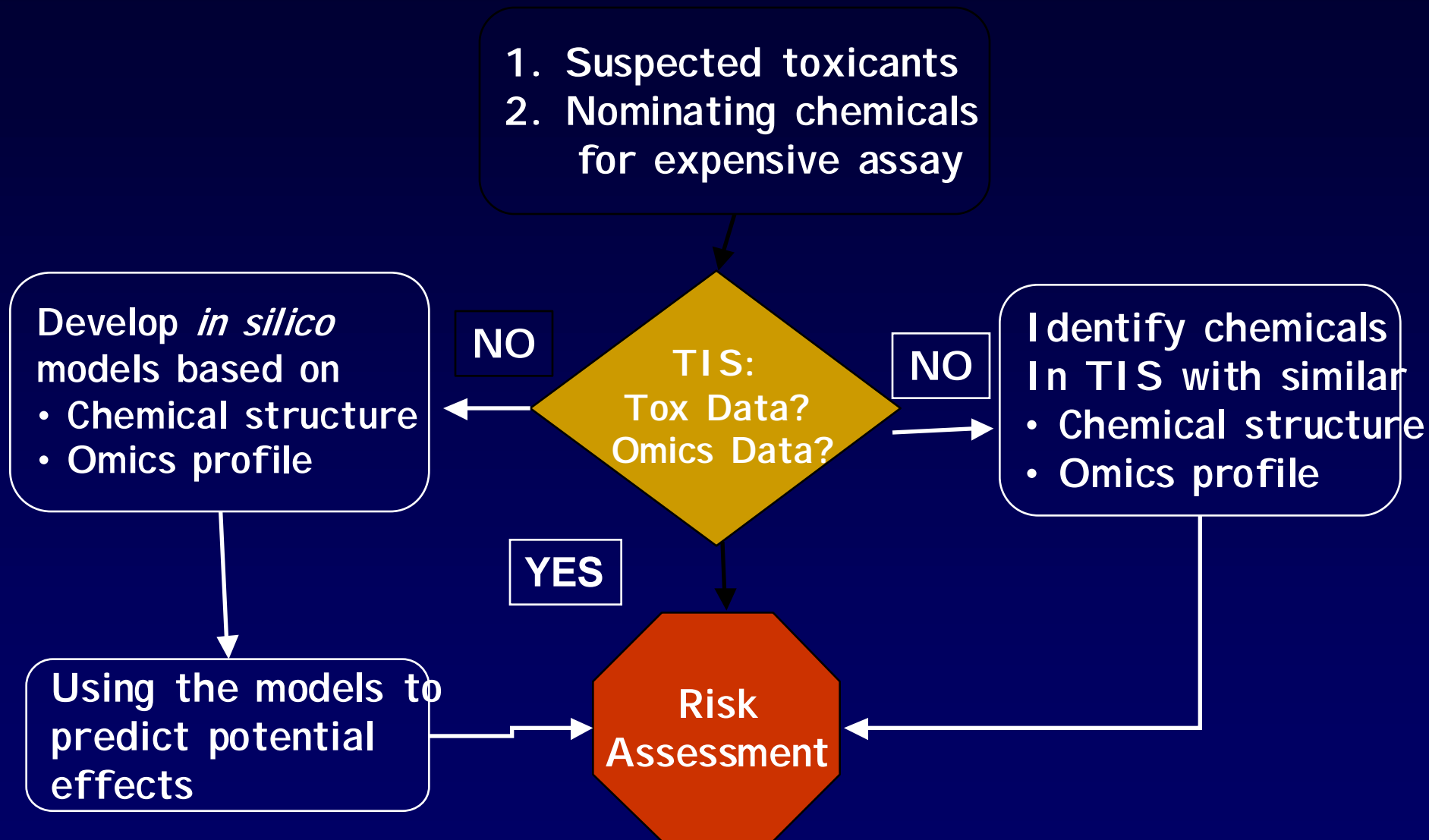


- Cross-linked between two libraries
- Link genomics data with proteomics data
- Link in-house data with public data

Toxicoinformatics Integrated System (TIS)



Toxicoinformatics and Risk Assessment



Systemic Experiments

- Acetaminophen
- Glitazone (NCTR, CDER, Merck)

Acknowledgment

- Jim Fuscoe
- Luke Ratnasinghe
- John Bowyer
- Rob Turesky
- Rick Edmondson
- Rick Beger
- Weida Tong
- Hong Fang
- Bob Delongchamp
- Jim Chen
- Yvonne Dragan